In Focus: Six degrees of Apodemus separation

Author

McCallum, Hamish

Published

2009

Journal Title

Journal of Animal Ecology

DOI

10.1111/j.1365-2656.2009.01572.x

In Focus

Six degrees of *Apodemus* separation.

Perkins, S. E., Cagnacci, F., Stradiotto, A., Arnoldi, D. & Hudson, P. J. (2009).A comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. *Journal of Animal Ecology* **xx**, xxxx-xxxx.

Social network analysis has been widely used to help understand the transmission of human diseases. Its application to wildlife disease is very much in its infancy, largely because of the difficulty of recording contacts between wild animals. Sarah Perkins and co-authors have constructed contact networks for yellow-necked mice (*Apodemus flavicollis*) in the Italian Alps, comparing networks derived from radio tracking and mark-recapture data. They found that the method producing the most informative data depended on population density. However, all networks had aggregated contact distributions, which is important for understanding disease transmission.

Transmission of any infectious disease involves a connection between an infected host and a susceptible host. The formerly esoteric topic of network theory, concerned with describing and understanding networks of connections between objects, has grasped the imagination of the public since the making of the film *Six Degrees of Separation*. It is therefore not surprising that much recent attention has been given to the applications of network theory to understanding the dynamics of infectious disease transmission (Meyers et al., 2005, Bansal et al., 2007, Lloyd-Smith et al., 2005, Keeling and Eames, 2005, Read et al., 2008, Kao et al., 2007). The universal

conclusion of all these studies is that the structure of the contact network has a profound effect on disease dynamics, leading to different conclusions from those based on the 'mean field' assumption, which is that the mean rate of contact in the population is sufficient to understand disease dynamics. Most of this work has concentrated on human diseases. Even in this case, collecting the fundamental data -- who has made contact with whom -- is not a straightforward task. The usual approach for humans involves individuals recalling or recording with whom they have made contact. This may involve tracing the source of infection for each diseased subject, tracing all individuals (whether infected or not) the subjects have encountered, or asking subjects to maintain a diary as they contact new individuals (Keeling and Eames, 2005). None of these is easy with human subjects. Collecting contact information for a wildlife population is substantially more difficult. In this *In Focus* article I highlight one of the first studies to construct individual level contact networks for a wildlife population in order to infer the potential for disease transmission (Perkins et al. 2009). There have been just a handful of earlier studies. Cross et al (2004) derived what seems to be the first empirical contact network relevant to disease transmission for bovine tuberculosis in African buffalo (*Syncerus caffer*). Whilst they used radio telemetry, as was done by Perkins et al, 'contact' was inferred by animals being within the same herd. Corner et al. (2003) actually followed transmission of bovine tuberculosis through a brushtail possum (*Trichosurus vulpecula*) population, but in captivity. More recently, Porphyre et al. (2008) produced a contact network for brushtail possums and bovine tuberculosis in the wild, inferring 'contact' either by animals' activity range or by capture at the same trap location within one month.

Perkins et al. (2009) used both radio telemetry and trapping data to investigate networks of contacts between yellow-necked mice *Apodemus flavicollis* in the Italian Alps. 'Contact' was inferred either by a mouse having a location within 15 m of another individual's location within a two-week radio tracking session, or being caught in the same or an adjacent trap (in a 15m spaced grid) within a two-week trapping session. The authors did not investigate any specific parasite, although they suggest that faecal-orally transmitted parasites might have transmission networks similar to these contact networks. They asked whether networks based on the two methods had similar network statistics. They did not, but the method appearing to produce the more informative data depended on the mouse population density, which differed substantially between the two years of the study. A question that remains open is whether this is indeed a function of the contact network changing with mouse density, or whether it is simply that a higher proportion of the total population was radio tracked in the year when the mouse population was lower. Because of the difficulties in inference, much of their analysis is based on qualitative descriptions, such as quantile-quantile plots. No doubt there is scope for devising more rigorous means of analysing data of this sort (for example, see Lusseau et al., 2008), but the authors have made a start and hopefully the paper will attract the attention of some keen mathematical statistics postgraduates.

Wildlife disease researchers will typically want to compare networks at different times or places with each other or to compare the connectedness of sexes or other classes of individual within a given network. Unfortunately, applying standard statistical approaches to these data is difficult. An edge connects two nodes (usually individual animals for disease networks) and multiple edges may depart from one

node. This means there are non-independence issues associated with inference based on many of these statistics. Randomisation methods provide a means to get around some of these issues. Statistical inference is particularly a problem for the relatively small networks most wildlife researchers need to work with and for the sorts of questions we are interested in answering. Compounding these problems is the fact that only a proportion of the animals interacting in a network will actually have been followed.

Despite its popularity, applying much of the published work in network theory to disease transmission is rather frustrating. Network papers use terms which are overlapping concepts in common English, although their technical definitions may describe quite different properties of networks. In Perkins et al., for example, the authors use two measures of 'centrality': 'closeness', which is a property of a single node (here, a mouse) defined as the shortest path between that node and all others in the network; and 'betweenness', again a property of a node, defined as the number of shortest paths between all other individuals that pass through the node in question (Wasserman and Faust 1994). A distinct concept is 'connectedness' which is a property of the network as a whole, defined as the proportion of all possible links that actually exist in network. It is evident that nodes with high betweenness and closeness are particularly important to disease transmission and that diseases will spread best across highly connected networks. However, which are the most relevant of these network properties for understanding the various aspects of disease transmission is unclear.

An important result Perkins *et al.* found was that the degree distribution in the networks (the frequency distribution of the number of contacts made by each mouse) was fitted better by a negative binomial distribution than a Poisson distribution. This was the case for networks based on either radio telemetry or mark-recapture, although the extent of overdispersion in the degree distribution varied both through time and between the two types of data. The significance of this result is that the basic reproductive number R_0 of a pathogen depends on the coefficient of variation of the degree distribution of the network *CV* according to the following equation:

$$R_0 = \rho_0 \left[1 + (CV)^2 \right] \quad (1)$$

where ρ_0 is a constant (May, 2006). R_0, the number of secondary cases per primary case when disease is first introduced is well-known to epidemiologists as key to understanding whether and how fast an infectious disease can invade a given population (Roberts, 2007, Cross et al., 2007, Heesterbeek, 2002). This result shows that a highly aggregated degree distribution will cause an infectious disease to spread more easily and rapidly than would be the case given the 'mean field' assumption made in most basic epidemiological models. In the specific case of a negative binomial distribution, using a result from Appendix (1) of Anderson and May (1978) equation (1) can be written as:

$$R_0 = \rho_0 \left[1 + \frac{m(k+1)}{k} \right] \quad (2)$$

Where *m* is the mean of the degree distribution and *k* is the parameter of the negative binomial distribution that inversely describes the extent of overdispersion.

So, where should investigations of wildlife disease contact networks go next? Perhaps the most crucial issue is to recognise that infectious disease transmission is a three-

stage process. First, there needs to be a contact between an infected and susceptible host such that a particular pathogen might be transmitted. This current study sheds important light on this stage, although whether either of the contact definitions used (having a location within 15 m of another individual's location within a two-week radio tracking session, or being caught in the same or an adjacent trap within a two-week trapping session) are relevant to any particular *Apodemus* pathogen remains an open question. As Cross et al. (2004) emphasise, the length of time over which contacts are integrated can have a substantial effect on network structure. Second, the contact actually has to result in the transfer of infectious propagules, Finally, those propagules need to develop successfully to produce infections capable of transmitting to another host. More sophisticated ways of logging contacts, such as proximity sensing radio collars (Ji et al., 2005) may improve resolution of epidemiologically relevant contacts, but the length or 'quality' of contacts might also influence the second and third stages of the transition process . Only following a novel pathogen or pathogen strain as it invades a naïve population is likely to shed light on these stages of the transmission process.

Finally, if we are to understand the implications of network structure on infectious disease dynamics in wildlife, there is a need to move from descriptive statistics to inference. A better understanding of the sampling properties of summary statistics for the relatively small, incomplete networks of relevance to wildlife disease is essential.

Acknowledgements

H. McCallum's work on wildlife disease networks is funded by an Australian Research Council Discovery Grant (DP0772644)

HAMISH McCALLUM

School of Zoology

University of Tasmania

Hobart, Tasmania 7001

Australia

References

Anderson, R. M. & May, R. M. (1978) Regulation and stability of host-parasite interactions. I. Regulatory processes. *Journal of Animal Ecology,* **47,** 219-247.

Bansal, S., Grenfell, B. T. & Meyers, L. A. (2007) When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface,* **4,** 879-891.

Corner, L. A. L., Pfeiffer, D. U. & Morris, R. S. (2003) Social-network analysis of Mycobacterium bovis transmission among captive brushtail possums (Trichosurus vulpecula). *Preventive Veterinary Medicine,* **59,** 147-167.

Cross, P., Johnson, P., Lloyd-Smith, J. & Getz, W. (2007) Utility of R_0 as a predictor of disease invasion in structured populations. *Journal of The Royal Society Interface,* **4,** 315-324.

Cross, P. C., Lloyd-Smith, J. O., Bowers, J. A., Hay, C. T., Hofmeyr, M. & Getz, W. M. (2004) Integrating association data and disease dynamics in a social ungulate: bovine tuberculosis in African buffalo in the Kruger National Park. *Annales Zoologici Fennici,* **41,** 879-892.

Heesterbeek, J. A. P. (2002) A Brief History of R_0 and a Recipe for its Calculation. *Acta Biotheoretica,* **50,** 189-204.

Ji, W. H., White, P. C. L. & Clout, M. N. (2005) Contact rates between possums revealed by proximity data loggers. *Journal of Applied Ecology,* **42,** 595-604.

Kao, R. R., Green, D. M., Johnson, J. & Kiss, I. Z. (2007) Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the UK. *Journal of the Royal Society Interface,* **4,** 907-916.

Keeling, M. J. & Eames, K. T. D. (2005) Networks and epidemic models. *Journal of the Royal Society Interface,* **2,** 295-307.

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. (2005) Superspreading and the effect of individual variation on disease emergence. *Nature,* **438,** 355-359.

Lusseau, D., Whitehead, H. & Gero, S. (2008) Incorporating uncertainty into the study of animal social networks. *Animal Behaviour,* **75,** 1809-1815.

May, R. M. (2006) Network structure and the biology of populations. *Trends in Ecology & Evolution,* **21,** 394-399.

Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M. & Brunham, R. C. (2005) Network theory and SARS: predicting outbreak diversity. *Journal of Theoretical Biology,* **232,** 71-81.

Perkins, S. E., F. Cagnacci, A. Stradiotto, D. Arnoldi, and P. J. Hudson. 2009. A comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. *Journal of Animal Ecology* **xx**:xxx-xxx.

Porphyre, T., Stevenson, M., Jackson, R. & McKenzie, J. (2008) Influence of contact heterogeneity on TB reproduction ratio R-0 in a free-living brushtail possum Trichosurus vulpecula population. *Veterinary Research,* **39,** 13.

Read, J. M., Eames, K. T. D. & Edmunds, W. J. (2008) Dynamic social networks and the implications for the spread of infectious disease. *Journal of the Royal Society Interface,* **5,** 1001-1007.

Roberts, M. G. (2007) The pluses and minuses of R-0. *Journal of the Royal Society Interface,* **4,** 949-961.

Wasserman, S. and K. Faust. 1994. *Social network analysis: Methods and applications.* Cambridge University Press, Cambridge.